

#2

OIPE

RAW SEQUENCE LISTING

DATE: 07/23/2001

PATENT APPLICATION: US/09/895,263

TIME: 13:27:57

Input Set : A:\PF140 NewSeq050801.txt

Output Set: N:\CRF3\07232001\I895263.raw

ENTERED

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: He, Wei-Wu et al.

8 (ii) TITLE OF INVENTION: Interleukin-1 Beta Converting Enzyme
9 Like Apoptosis Protease 3 and 4

11 (iii) NUMBER OF SEQUENCES: 12

13 (iv) CORRESPONDENCE ADDRESS:

14 (A) ADDRESSEE: Human Genome Sciences, Inc.

15 (B) STREET: 9410 Key West Ave.

16 (C) CITY: Rockville

17 (D) STATE: MD

18 (E) COUNTRY: USA

19 (F) ZIP: 20850

21 (v) COMPUTER READABLE FORM:

22 (A) MEDIUM TYPE: Floppy disk

23 (B) COMPUTER: IBM PC compatible

24 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

27 (vi) CURRENT APPLICATION DATA:

C--> 28 (A) APPLICATION NUMBER: US/09/895,263

C--> 29 (B) FILING DATE: 02-Jul-2001

30 (C) CLASSIFICATION:

32 (vii) PRIOR APPLICATION DATA:

33 (A) APPLICATION NUMBER:

34 (B) FILING DATE:

36 (viii) ATTORNEY/AGENT INFORMATION:

37 (A) NAME: Jonathan L. Klein

38 (B) REGISTRATION NUMBER: 41,119

39 (C) REFERENCE/DOCKET NUMBER: PF140

41 (ix) TELECOMMUNICATION INFORMATION:

42 (A) TELEPHONE: 301-251-6015

43 (B) TELEFAX: 301-309-8439

46 (2) INFORMATION FOR SEQ ID NO: 1:

48 (i) SEQUENCE CHARACTERISTICS:

49 (A) LENGTH: 1369 base pairs

50 (B) TYPE: nucleic acid

51 (C) STRANDEDNESS: single

52 (D) TOPOLOGY: linear

54 (ii) MOLECULE TYPE: DNA (genomic)

57 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```

59 GCACGAGAAA CTTTGCTGTG CGCGTTCTCC CGCGCGCGGG CTCAACTTTG TAGAGCGAGG 60
61 GGCCAACCTG GCAGAGCGCG CGGCCAGCTT TGCAGAGAGC GCCCTCCAGG GACTATGCGT 120
63 GCGGGGACAC GGGTCGCTTT GGGCTCTTCC ACCCTGCGG AGCGCACTAC CCCGAGCCAG 180
65 GGGCGGTGCA AGCCCCGCCC GGCCCTACCC AGGGCGGCTC CTCCCTCCGC AGCGCCGAGA 240
67 CTTTTAGTTT CGCTTTGCTT AAAGGGGCCC CAGACCCTTG CTGCGGAGCG ACGGAGAGAG 300
69 ACTGTGCCAG TCCCAGCCGC CCTACCGCCG TGGGAACGAT GGCAGATGAT CAGGGCTGTA 360
71 TTGAAGAGCA GGGGGTTGAG GATTTCAGCAA ATGAAGATTC AGTGGATGCT AAGCCAGACC 420

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/895,263

DATE: 07/23/2001

TIME: 13:27:57

Input Set : A:\PF140 NewSeq050801.txt

Output Set: N:\CRF3\07232001\I895263.raw

```

73 GGTCTCGTTC TGTACCGTCC CTCTTCAGTA AGAAGAAGAA AAATGTCACC ATGCGATCCA 480
75 TCAAGACCAC CCGGGACCGA GTGCCTACAT ATCAGTACAA CATGAATTTT GAAAAGCTGG 540
77 GCAAATGCAT CATAATAAAC AACAAGAAGT TGTATAAAGT GACAGGTATG GGCGTTCGAA 600
79 ACGGAACAGA CAAAGATGCC GAGGCGCTCT TCAAGTGCTT CCGAAGCCTG GGTTTTGACG 660
81 TGATTGTCTA TAATGACTGC TCTTGTGCCA AGATGCAAGA TCTGCTTAAA AAAGCTTCTG 720
83 AAGAGGACCA TACAAATGCC GCCTGCTTCG CCTGCATCCT CTTAAGCCAT GGAGAAGAAA 780
85 ATGTAATTTA TGGGAAAGAT GGTGTCACAC CAATAAAGGA TTTGACAGCC CACTTTAGGG 840
87 GGGATAGATG CAAAACCTTT TTAGAGAAAC CCAAACCTTT CTTCAATCAG GCTTGCCGAG 900
89 GGACCGAGCT TGATGATGCC ATCCAGGCCG ACTCGGGGCC CATCAATGAC ACAGATGCTA 960
91 ATCCTCGATA CAAGATCCCA GTGGAAGCTG ACTTCTCTT CGCCTATTCC ACGGTTCCAG 1020
93 GCTATTACTC GTGGAGGAGC CCAGGAAGAG GCTCCTGGTT TGTGCAAGCC CTCTGCTCCA 1080
95 TCCTGGAGGA GCACGGAAAA GAGCTGGAAA TCATGCAAAT CCTCACCAGG GTGAATGACA 1140
97 GAGTTGCCAG GCACTTTGAG TCTCAGTCTG ATGACCCACA CTTCCATGAG AAGAAGCAGA 1200
99 TCCCCTGTGT GGTCTCCATG CTCACCAAGG AACTCTACTT CAGTCAATAG CCATATCAGG 1260
101 GGTACATTCT AGCTGAGAAG CAATGGGTCA CTCATTAATG AATCACATTT TTTTATGCTC 1320
103 TTGAAATATT CAGAAATTCT CCAGGATTTT AATTTTCAGGA AAATGTATT 1369

```

106 (2) INFORMATION FOR SEQ ID NO: 2:

108 (i) SEQUENCE CHARACTERISTICS:

109 (A) LENGTH: 303 amino acids

110 (B) TYPE: amino acid

111 (C) STRANDEDNESS: single

112 (D) TOPOLOGY: linear

114 (ii) MOLECULE TYPE: protein

117 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

119 Met Ala Asp Asp Gln Gly Cys Ile Glu Glu Gln Gly Val Glu Asp Ser
120 1 5 10 15
122 Ala Asn Glu Asp Ser Val Asp Ala Lys Pro Asp Arg Ser Ser Phe Val
123 20 25 30
125 Pro Ser Leu Phe Ser Lys Lys Lys Lys Asn Val Thr Met Arg Ser Ile
126 35 40 45
128 Lys Thr Thr Arg Asp Arg Val Pro Thr Tyr Gln Tyr Asn Met Asn Phe
129 50 55 60
131 Glu Lys Leu Gly Lys Cys Ile Ile Ile Asn Asn Lys Asn Phe Asp Lys
132 65 70 75 80
134 Val Thr Gly Met Gly Val Arg Asn Gly Thr Asp Lys Asp Ala Glu Ala
135 85 90 95
137 Leu Phe Lys Cys Phe Arg Ser Leu Gly Phe Asp Val Ile Val Tyr Asn
138 100 105 110
140 Asp Cys Ser Cys Ala Lys Met Gln Asp Leu Leu Lys Lys Ala Ser Glu
141 115 120 125
143 Glu Asp His Thr Asn Ala Ala Cys Phe Ala Cys Ile Leu Leu Ser His
144 130 135 140
146 Gly Glu Glu Asn Val Ile Tyr Gly Lys Asp Gly Val Thr Pro Ile Lys
147 145 150 155 160
149 Asp Leu Thr Ala His Phe Arg Gly Asp Arg Cys Lys Thr Leu Leu Glu
150 165 170 175
152 Lys Pro Lys Leu Phe Phe Ile Gln Ala Cys Arg Gly Thr Glu Leu Asp
153 180 185 190
155 Asp Ala Ile Gln Ala Asp Ser Gly Pro Ile Asn Asp Thr Asp Ala Asn

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/895,263

DATE: 07/23/2001

TIME: 13:27:57

Input Set : A:\PF140 NewSeq050801.txt

Output Set: N:\CRF3\07232001\I895263.raw

```

156          195          200          205
158    Pro Arg Tyr Lys Ile Pro Val Glu Ala Asp Phe Leu Phe Ala Tyr Ser
159          210          215          220
161    Thr Val Pro Gly Tyr Tyr Ser Trp Arg Ser Pro Gly Arg Gly Ser Trp
162    225          230          235          240
164    Phe Val Gln Ala Leu Cys Ser Ile Leu Glu Glu His Gly Lys Glu Leu
165          245          250          255
167    Glu Ile Met Gln Ile Leu Thr Arg Val Asn Asp Arg Val Ala Arg His
168          260          265          270
170    Phe Glu Ser Gln Ser Asp Asp Pro His Phe His Glu Lys Lys Gln Ile
171          275          280          285
173    Pro Cys Val Val Ser Met Leu Thr Lys Glu Leu Tyr Phe Ser Gln
174    290          295          300

176 (2) INFORMATION FOR SEQ ID NO: 3:
178   (i) SEQUENCE CHARACTERISTICS:
179       (A) LENGTH: 1159 base pairs
180       (B) TYPE: nucleic acid
181       (C) STRANDEDNESS: single
182       (D) TOPOLOGY: linear
184   (ii) MOLECULE TYPE: DNA (genomic)
187   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
189 GCACGAGCGG ATGGGTGCTA TTGTGAGGCG GTTGTAGAAG AGTTTCGTGA GTGCTCGCAG      60
191 CTCATACCTG TGGCTGTGTA TCCGTGGCCA CAGCTGGTTG GCGTCGCCTT GAAATCCAG      120
193 GCCGTGAGGA GTTAGCGAGC CCTGCTCACA CTCGGCGCTC TGGTTTTTCGG TGGGTGTGCC      180
195 CTGCACCTGC CTCTTCCCGC ATTCTCATTA ATAAAGGTAT CCATGGAGAA CACTGAAAAC      240
197 TCAGTGGAAT CAAAATCCAT TAAAAATTTG GAACCAAAGA TCATACATGG AAGCGAATCA      300
199 ATGGACTCTG GAATATCCCT GGACAACAGT TATAAAATGG ATTATCCTGA GATGGGTTTA      360
201 TGTATAATAA TTAATAATAA GAATTTTCAT AAAAGCACTG GAATGACATC TCGGTCTGGT      420
203 ACAGATGTCT ATGCAGCAAA CCTCAGGGAA ACATTCAGAA ACTTGAAATA TGAAGTCAGG      480
205 AATAAAAATG ATCTTACACG TGAAGAAATT GTGGAATTGA TCGGTGATGT TTCTAAAGAA      540
207 GATCACAGCA AAAGGAGCAG TTTTGTGTTG GTGCTTCTGA GCCATGGTGA AGAAGGAATA      600
209 ATTTTGTGAA CAAATGGACC TGTGACCTG AAAAAAATAA CAAACTTTTT CAGAGGGGAT      660
211 CGTTGTAGAA GTCTAACTGG AAAACCCAAA CTTTTCATTA TTCAGGCCTG CCGTGGTACA      720
213 GAACCTGGACT GTGGCATTGA GACAGACAGT GGTGTTGATG ATGACATGGC GTGTCATAAA      780
215 ATACCAGTGG AGGCCGACTT CTTGTATGCA TACTCCACAG CACCTGGTTA TTATTCTTGG      840
217 CGAAATTCAA AGGATGGCTC CTGGTTCATC CAGTCGCTTT GTGCCATGCT GAAACAGTAT      900
219 GCCGACAAGC TTGAATTTAT GCACATTCTT ACCCGGGTTA ACCGAAAGGT GGCAACAGAA      960
221 TTTGAGTCCT TTTCTTTTGA CGCTACTTTT CATGCAAAGA AACAGATTCC ATGTATTGTT     1020
223 TCCATGCTCA CAAAAGAACT CTATTTTAT CACTAAAGAA ATGGTTGGTT GGTGGTTTTT     1080
225 TTTAGTTTGT ATGCCAAGTG AGAAGATGGT ATATTGGGT ACTGTATTTC CCTCTCATTG     1140
227 GGGACCTACT CTCATGCTG                                     1159

229 (2) INFORMATION FOR SEQ ID NO: 4:
231   (i) SEQUENCE CHARACTERISTICS:
232       (A) LENGTH: 277 amino acids
233       (B) TYPE: amino acid
234       (C) STRANDEDNESS: single
235       (D) TOPOLOGY: linear
237   (ii) MOLECULE TYPE: protein
240   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/895,263

DATE: 07/23/2001

TIME: 13:27:57

Input Set : A:\PF140 NewSeq050801.txt

Output Set: N:\CRF3\07232001\I895263.raw

```

242 Met Glu Asn Thr Glu Asn Ser Val Asp Ser Lys Ser Ile Lys Asn Leu
243 1 5 10 15
245 Glu Pro Lys Ile Ile His Gly Ser Glu Ser Met Asp Ser Gly Ile Ser
246 20 25 30
248 Leu Asp Asn Ser Tyr Lys Met Asp Tyr Pro Glu Met Gly Leu Cys Ile
249 35 40 45
251 Ile Ile Asn Asn Lys Asn Phe His Lys Ser Thr Gly Met Thr Ser Arg
252 50 55 60
254 Ser Gly Thr Asp Val Asp Ala Ala Asn Leu Arg Glu Thr Phe Arg Asn
255 65 70 75 80
257 Leu Lys Tyr Glu Val Arg Asn Lys Asn Asp Leu Thr Arg Glu Glu Ile
258 85 90 95
260 Val Glu Leu Met Arg Asp Val Ser Lys Glu Asp His Ser Lys Arg Ser
261 100 105 110
263 Ser Phe Val Cys Val Leu Leu Ser His Gly Glu Glu Gly Ile Ile Phe
264 115 120 125
266 Gly Thr Asn Gly Pro Val Asp Leu Lys Lys Ile Thr Asn Phe Phe Arg
267 130 135 140
269 Gly Asp Arg Cys Arg Ser Leu Thr Gly Lys Pro Lys Leu Phe Ile Ile
270 145 150 155 160
272 Gln Ala Cys Arg Gly Thr Glu Leu Asp Cys Gly Ile Glu Thr Asp Ser
273 165 170 175
275 Gly Val Asp Asp Asp Met Ala Cys His Lys Ile Pro Val Glu Ala Asp
276 180 185 190
278 Phe Leu Tyr Ala Tyr Ser Thr Ala Pro Gly Tyr Tyr Ser Trp Arg Asn
279 195 200 205
281 Ser Lys Asp Gly Ser Trp Phe Ile Gln Ser Leu Cys Ala Met Leu Lys
282 210 215 220
284 Gln Tyr Ala Asp Lys Leu Glu Phe Met His Ile Leu Thr Arg Val Asn
285 225 230 235 240
287 Arg Lys Val Ala Thr Glu Phe Glu Ser Phe Ser Phe Asp Ala Thr Phe
288 245 250 255
290 His Ala Lys Lys Gln Ile Pro Cys Ile Val Ser Met Leu Thr Lys Glu
291 260 265 270
293 Leu Tyr Phe Tyr His
294 275

```

296 (2) INFORMATION FOR SEQ ID NO: 5:

298 (i) SEQUENCE CHARACTERISTICS:

299 (A) LENGTH: 31 base pairs

300 (B) TYPE: nucleic acid

301 (C) STRANDEDNESS: single

302 (D) TOPOLOGY: linear

304 (ii) MOLECULE TYPE: DNA (genomic)

307 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

309 GATCGGATCC ATGCGTGCGG GGACACGGGT C

311 (2) INFORMATION FOR SEQ ID NO: 6:

313 (i) SEQUENCE CHARACTERISTICS:

314 (A) LENGTH: 31 base pairs

315 (B) TYPE: nucleic acid

31

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/895,263

DATE: 07/23/2001

TIME: 13:27:57

Input Set : A:\PF140 NewSeq050801.txt

Output Set: N:\CRF3\07232001\I895263.raw

```

316          (C) STRANDEDNESS: single
317          (D) TOPOLOGY: linear
319      (ii) MOLECULE TYPE: DNA (genomic)
322      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
324 GTACTCTAGA TCATTACCC TGGTGGAGGA T                               31
326 (2) INFORMATION FOR SEQ ID NO: 7:
328      (i) SEQUENCE CHARACTERISTICS:
329          (A) LENGTH: 31 base pairs
330          (B) TYPE: nucleic acid
331          (C) STRANDEDNESS: single
332          (D) TOPOLOGY: linear
334      (ii) MOLECULE TYPE: DNA (genomic)
337      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
339 GATCGGATCC ATGGAGAACA CTGAAACTC A                               31
341 (2) INFORMATION FOR SEQ ID NO: 8:
343      (i) SEQUENCE CHARACTERISTICS:
344          (A) LENGTH: 31 base pairs
345          (B) TYPE: nucleic acid
346          (C) STRANDEDNESS: single
347          (D) TOPOLOGY: linear
349      (ii) MOLECULE TYPE: DNA (genomic)
352      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
354 GTACTCTAGA TTAGTGATAA AAATAGAGTT C                               31
356 (2) INFORMATION FOR SEQ ID NO: 9:
358      (i) SEQUENCE CHARACTERISTICS:
359          (A) LENGTH: 22 base pairs
360          (B) TYPE: nucleic acid
361          (C) STRANDEDNESS: single
362          (D) TOPOLOGY: linear
364      (ii) MOLECULE TYPE: DNA (genomic)
367      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
369 GACTATGCGT GCGGGGACAC GG                                         22
371 (2) INFORMATION FOR SEQ ID NO: 10:
373      (i) SEQUENCE CHARACTERISTICS:
374          (A) LENGTH: 53 base pairs
375          (B) TYPE: nucleic acid
376          (C) STRANDEDNESS: single
377          (D) TOPOLOGY: linear
379      (ii) MOLECULE TYPE: DNA (genomic)
382      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
384 AATCAAGCGT AGTCTGGGAC GTCGTATGGG TATTCACCCT GGTGGAGGAT TTG       53
386 (2) INFORMATION FOR SEQ ID NO: 11:
388      (i) SEQUENCE CHARACTERISTICS:
389          (A) LENGTH: 21 base pairs
390          (B) TYPE: nucleic acid
391          (C) STRANDEDNESS: single
392          (D) TOPOLOGY: linear
394      (ii) MOLECULE TYPE: DNA (genomic)
397      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/895,263

DATE: 07/23/2001

TIME: 13:27:58

Input Set : A:\PF140 NewSeq050801.txt

Output Set: N:\CRF3\07232001\I895263.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

L:418 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:12 —